

Peer-review report of

Wysocki, A., McCarthy, I., van Bork, R., Cramer, A.O.J., & Rhemtulla, M. (2025). Cross-lagged panel networks. *advances.in/psychology*, 2, e739621. <https://doi.org/10.56296/aip00037>

Round 1

Dear Authors,

Thank you for submitting your manuscript, titled "Cross-Lagged Network Models" to our special issue on network psychometrics. Your work has undergone a thorough review process, and I have received detailed feedback from a reviewer with expertise in network psychometrics.

The reviewer has provided comprehensive comments, highlighting several areas where further clarification, expansion, or modification could enhance the manuscript's quality and contribution to the field. The feedback is constructive and aimed at refining your work to ensure its impact and accessibility to the intended audience. Below, I summarize the key points from the review that require attention:

- 1. Clarification on Sparse Networks (Page 7):** The term "sparsest" used to describe networks generated by l_1 -regularized regression techniques may be misleading. A more accurate descriptor might be "sparse," considering that non-regularized methods can result in even sparser networks.
- 2. Expansion of Software Packages (Footnote on Page 7):** Including additional software packages like {mgm}, {BGGM}, and {GGMnonreg} would offer a more comprehensive overview of tools available for estimating partial correlation networks.
- 3. Detail on Stationary Assumption (Pages 9 and 11):** Further explanation is requested on why the stationary assumption can be disregarded in your approach, especially given the use of autoregressive models.
- 4. Guidance on Collapsing Nodes (Pages 13-14):** The section on collapsing nodes based on semantic similarity lacks clear direction. Suggestions include discussing statistical methods or natural language processing techniques for more objective decision-making.
- 5. Clarification on Selection of λ (Page 16):** The process for selecting λ using 10-fold cross-validation needs clearer explanation, particularly regarding how the final λ value is determined from the cross-validation process.
- 6. Interpretation of Individual Paths (Page 29):** Insights into how the provided interpretations apply to the sample or subsets thereof would be valuable, including any potential for identifying specific trajectories for individual students or groups.
- 7. Discussion on Predictive Power at Variable-Level (Referenced Study):** Highlighting the findings from personality research on the enhanced predictive power of individual items over their aggregates could contextualize your results.

8. Explanation of RI-CLPM (Page 38): Additional details on how a single factor model representing the random-intercept is applied and its feasibility within the current framework would benefit the reader's understanding.

Overall, the reviewer acknowledges the manuscript's readability and its potential impact on the field. The suggested revisions are intended to further polish and strengthen your manuscript.

We invite you to revise your manuscript in line with the reviewer's feedback and resubmit it for consideration. Please provide a detailed response to each point raised, outlining the changes made in the manuscript accordingly. If you disagree with any suggestions, kindly provide your rationale.

Your revised manuscript will undergo a follow-up review to assess the incorporation of the suggested changes. We appreciate your commitment to enhancing your work and look forward to receiving your revised submission.

Should you have any questions or require clarification on the review comments, please do not hesitate to contact me.

Thank you for considering our new journal as a venue for your research. We believe that your manuscript, upon revision, will make a significant contribution to the field.

Best regards,

Hudson Golino

Reviewer 1:

In this manuscript, the authors propose an algorithm for estimating a model termed the cross-lagged panel network (CLPN), which is closely related to the cross-lagged panel model (CLPM). I have been familiar with this manuscript for some time, as I understand a pre-print has been available already for several years. The proposed CLPN method has also been applied already in several empirical studies (e.g., Funkhouser et al., 2021; Liu et al., 2024; Zainal & Newman, 2023a), although it seems some of these studies are based on an outdated pre-print and do not apply the full two-step procedure proposed in the manuscript. Given that this manuscript is already shaping up applied research, I think it is very important that it will soon be published in a crystallized form. The proposed method also fills a clear gap in modeling Gaussian variables using network models, as currently commonly used methods can only be applied to single-measurement (cross-sectional) data (Epskamp, Haslbeck, & Borkulo, 2022), panel data with at least three waves of data (Epskamp, 2020), or high intensity time-series data with at least 20 measurements per person (Epskamp, Waldorp, et al., 2018). However, I do have some concerns with the manuscript, especially since it is already shaping up empirical research. I have some methodological concerns that I think are very important to address to support applied research that uses the CLPN method. Notably, I think that (1) the method should be implemented in accessible open-source software to minimize the risk of applied researchers implementing the method wrongly, (2) the method should

be supported with simulation studies to gain insight in the sensitivity and specificity of the proposed estimation algorithm, and (3) clear guidelines for assessing the accuracy and stability of results should be provided (ideally also using accessible software). In addition to these methodological concerns, I also have some conceptual concerns that I found important to mention. I recognize, however, that these concerns might be more academic in nature and some might be better placed in open academic discourse (e.g., through a commentary) instead of through the closed review process (in fact, I already co-authored one paper touching on some of these issues with the CLPN; Freichel et al., 2024).

Methodological concerns

Software implementation and availability

The method is currently, unlike many other methods in network psychometrics, not implemented in accessible open-source software. This will make reproducing the method harder for applied researchers and more prone to error. In addition, not having a standardized open source software implementation would make it harder for the method to be used in bootstrapped accuracy and stability checks that are now commonly used in network psychometrics (Epskamp, Borsboom, & Fried, 2018; Fried et al., 2022).¹ To this end, I strongly encourage the authors to implement 1The CLPN can be used in the bootnet R package already, see for example <https://github.com/SachaEpskamp/bootnet/issues/116>. However, note that this implementation does not use the second SEM fit step proposed in the manuscript.

The method in an openly available R package. R scripts with functions on the Open Science Framework are not sufficient for this purpose as these are hard to run, not well documented, and can break down in future R versions. If the authors wish for assistance with this I am very happy to help out (e.g., the method could also be implemented in the bootnet package and possibly also the first step in the mgm package).

Methodological evidence

Currently, the manuscript proposes a method, but does not provide any methodological evidence that the method will work. I am very happy to see the second SEM fit step proposed by the authors, as the cross-validation lasso regularization method is prone to false positives (Isvoranu & Epskamp, 2022b). I expect that the proposed method is quite conservative, but it would be good to see if this is correct and that the method has acceptable sensitivity and specificity in a large-scale simulation study (see for some pointers Hoekstra et al., 2023). It might also be good to critically evaluate other choices that could be made in the model selection algorithm, such as using EBIC instead of CV for LASSO model selection or by performing stepwise search algorithms in the second SEM-based step (Blanken et al., 2022; Isvoranu & Epskamp, 2022b). For example, I wonder if the first step using LASSO regularization is really needed, or if the parameter pruning performed

in SEM in the second step alone would be sufficient to estimate the network structure, given that such a thresholding/pruning algorithm is a common method for estimating non-regularized network models (Blanken et al., 2022; Isvoranu & Epskamp, 2022b).

Stability & accuracy

Some form of stability and/or accuracy analysis is usually performed when estimating psychometric network models (Epskamp, Borsboom, & Fried, 2018; Fried et al., 2022). This is very important, as network models can be quite unstable especially in smaller sample sizes. As such, there is a risk that results in unstable networks do not replicate well in new datasets. It would be useful if the manuscript details a possible set of accuracy analyses that can be performed after estimating a CLPN network. The manuscript now uses bootstrapping to estimate intervals on the in- and out-prediction estimates, but these bootstrap samples could also be used to estimate the stability of the network itself (this is mentioned in footnote 4 but could be shown in the paper).

Conceptual concerns

Embedding in the literature

The manuscript could be better embedded in the literature. I understand that a likely cause of this is the long writing duration this manuscript went through, but I do think that it would be good to include up-to-date references and discussions. For example, many of the topics touched on in this manuscript are now covered elaborately in the textbook on network psychometrics (Isvoranu et al., 2022). More importantly, the manuscript makes it seem that there currently are no network modeling techniques for panel data. This is not correct, the panelGVAR model (Epskamp, 2020) has already been available in open-source software for several years and has already been applied in several empirical studies (e.g., Freichel, Pfirrmann, et al., 2023; Freichel, Skjerdingsstad, et al., 2023; O'Driscoll et al., 2022; Schade et al., 2021; Zainal & Newman, 2023b). This approach also overcomes some of the challenges mentioned in the manuscript, such as separating within- and between-person effects and being able to incorporate measurement error.

Beyond this approach, Blanken et al. (2019) proposed network intervention analysis which can be applied to panel data, Conlin et al. (2022) proposed a method that is similar to the CLPN method discussed in the manuscript, and Koelen et al. (2022) use other alternative options for modeling two waves of data using networks. It would be good to update the manuscript with recent relevant literature and describe how the CLPN method differs from other network modeling approaches to panel data (see also Freichel et al., 2024). CLPN vs CLPM

The manuscript introduces the CLPN as an alternative to the CLPM. However, I don't really see where the CLPN differs structurally from the CLPM except for the estimation algorithm used. Isn't CLPN simply an exploratory way of estimating the

CLPM coupled with showing the obtained regression coefficients as a network? The manuscript presents the CLPM as a method for latent variables, but as far as I know the CLPM is usually applied to observed variables. To this end, is the CLPN really that different from the CLPM to warrant a different name? Other terms used for commonly used network models, such as the Ising model, Gaussian graphical model, and vector-autoregressive model, are used regardless of the estimation method used and regardless of whether or not the results are visualized and analyzed as a network structure. However, I do recognize that the CLPN name has already been used in the literature in several empirical papers (albeit it usually refers to only step 1 of the two-step procedure proposed), so I also see a good argument for keeping the name of this method.

Separating within- and between-person effects

Related to the above, a clear downside of the proposed CLPN method is that, like the CLPM, it cannot separate within-person effects (relationships between deviations from the mean) from between-person effects (relationships between stable individual differences; Hamaker et al., 2015). This means that a temporal effect can be found between two variables over time merely because there is a correlation between the stable averages of people over time (Epskamp, Hoekstra, et al., 2022). As such, this can indicate that a directed arrow in the estimated temporal effect doesn't equate to a temporal effect at all. This is very important reason for other network approaches on longitudinal data to separate within- and between-person effects (Burger et al., 2022; Epskamp, Waldorp, et al., 2018; Epskamp, 2020). This downside is now mentioned in the discussion already, but I think it should be elaborated on and be made more clear already in the introduction of the manuscript.

Cross-sectional network analysis

It is not really clear to me why Figure 2 is included. The proposed CLPN method does not make use of partial correlation networks and isn't really related to these (especially since the CLPN method does not include a contemporaneous network). Is the intent to show that such a cross-sectional network can be estimated per wave of data? What algorithm should researchers use then? What estimation method and what software are actually used to produce Figure 8? The text merely states a "non-regularized" method was applied, but there are several non-regularized methods to choose from, such as ggmModSelect, GGMnonreg, or several model selection routines in psychometrics (Blanken et al., 2022; Isvoranu & Epskamp, 2022b). An algorithm that is similar to the CLPN algorithm could be conceived for cross-sectional partial-correlation networks (LASSO with CV followed by maximum-likelihood based pruning in psychometrics) but should then be described and evaluated as well. If the authors propose to use a different estimation method for cross-sectional waves of the data (e.g., EBICglasso or ggmModSelect) it would be good to argue why a different estimation algorithm is warranted for cross-sectional networks compared to temporal CLPN networks. Perhaps the easiest and clearest would be simply to

remove Figure 2 and only describe cross-sectional networks briefly conceptually (with appropriate references).

Other concerns

Below, I list some other more minor concerns in no particular order.

- The wording of the introduction on pages 3 and 4 is quite strong in terms of causal language. However, the next section immediately takes back this strong causal interpretation: “Because of the challenge with making causal inference from cross-lagged parameters (Hamaker, Kuiper, & Grasman, 2015), we interpret cross-lagged paths in terms of prediction and not in causal terms.” I don’t disagree with the less strong causal tone, but I do find this quite a bit of a discrepancy in wording in a span of two pages.
- Page 27: The fit measures presented are not really interpretable, as exploratory search in the same dataset was used to obtain the network structure fitted.
- Page 9 reads that “The VAR approach can also be used on panel data, but it requires the assumption of stationary across the course of the time points (i.e., equal means and covariances across all timepoints) and returns a single network to represent all timepoints,” however, this is not supported by any references. The provided reference (Epskamp, Waldorp, et al., 2018) instead only focuses on datasets with at least about 20 time-points per person.
- The opening sentences on page 3 seem to presume that the average reader will be familiar with CLPMs, but I am not sure if that will be the case.
- Figure 1: Perhaps the covariances could be formatted a bit better, the lines are now thinner than the other lines and the arrowheads look different.
- The network figures in this paper are plotted using the qgraph default color theme (green and red). However, recently it is more common to use the colorblind friendly blue/red theme obtained with `theme = "colorblind"` (Isvoranu & Epskamp, 2022a).
- Page 8 states “These cross-network connections generate hypotheses about the pathways by which the two constructs could influence each other over time.” Perhaps some relevant work here could be cited (e.g., Epskamp, Waldorp, et al., 2018; Epskamp, Haslbeck, & Borkulo, 2022).
- Page 8: It would be fair to cite also the work by Bringmann et al. (2013) in the description of VAR models.
- Page 16: the abbreviation OLS is used before it is introduced. The abbreviation CV is also not introduced.
- What α level is used in the step 2 pruning step?
- Shouldn’t the error bars in Figures 6, 7, and 8 be truncated at zero?

Signed,

Sacha Epskamp

References

- Blanken, T. F., Isvoranu, A.-M., & Epskamp, S. (2022). Estimating network structures using model selection. In *Network psychometrics with R* (pp. 111–132). Routledge.
- Blanken, T. F., Van Der Zweerde, T., Van Straten, A., Van Someren, E. J., Borsboom, D., & Lancee, J. (2019). Introducing network intervention analysis to investigate sequential, symptom-specific treatment effects: A demonstration in co-occurring insomnia and depression. *Psychotherapy and Psychosomatics*, 88(1), 52–54.
- Bringmann, L. F., Vissers, N., Wichers, M., Geschwind, N., Kuppens, P., Peeters, F., Borsboom, D., & Tuerlinckx, F. (2013). A network approach to psychopathology: New insights into clinical longitudinal data. *PLoS One*, 8(4), e60188.
- Burger, J., Hoekstra, R. H. A., Mansueto, A. C., & Epskamp, S. (2022). Network estimation from time series and panel data. In *Network psychometrics with R* (pp. 169–192). Routledge.
- Conlin, W. E., Hoffman, M., Steinley, D., & Sher, K. J. (2022). Cross-sectional and longitudinal AUD symptom networks: They tell different stories. *Addictive Behaviors*, 131, 107333.
- Epskamp, S., Waldorp, L. J., Möttus, R., & Borsboom, D. (2018). The Gaussian graphical model in cross-sectional and time-series data. *Multivariate Behavioral Research*, 53(4), 453–480.
- Epskamp, S. (2020). Psychometric network models from time-series and panel data. *Psychometrika*, 85(1), 206–231.
- Epskamp, S., Borsboom, D., & Fried, E. I. (2018). Estimating psychological networks and their accuracy: A tutorial paper. *Behavior Research Methods*, 50, 195–212.
- Epskamp, S., Haslbeck, J. M. B., & Borkulo, A.-M. I. C. D. V. (2022). Pairwise Markov random fields. In *Network psychometrics with R* (pp. 93–110). Routledge.
- Epskamp, S., Hoekstra, R. H. A., Burger, J., & Waldorp, L. J. (2022). Longitudinal design choices: Relating data to analysis. In *Network psychometrics with R* (pp. 157–168). Routledge.
- Freichel, R., Pfirrmann, J., Cousjin, J., de Jong, P., Franken, I., Banaschewski, T., Bokde, A. L., Desrivères, S., Flor, H., Grigis, A., et al. (2023). Drinking motives, personality traits and life stressors—Identifying pathways to harmful alcohol use in adolescence using a panel network approach. *Addiction*, 118(10), 1908–1919.
- Freichel, R., Skjerdingsstad, N., Mansueto, A. C., Epskamp, S., Hoffart, A., Johnson, S. U., & Ebrahimi, O. V. (2023). Use of substances to cope predicts posttraumatic stress disorder symptom persistence: Investigating patterns of interactions between

symptoms and its maintaining mechanisms. *Psychological Trauma: Theory, Research, Practice, and Policy*.

Freichel, R., Veer, I., Wiers, R., McNally, R. J., & Epskamp, S. (2024). Cross-lagged panel models for studying psychopathology: A comparative overview of structural equation and panel network approaches. <https://doi.org/10.31234/osf.io/b94qt>

Fried, E. I., Epskamp, S., Veenman, M., & van Borkulo, C. D. (2022). Network stability, comparison, and replicability. In *Network psychometrics with R* (pp. 133–153). Routledge.

Funkhouser, C. J., Chacko, A. A., Correa, K. A., Kaiser, A. J., & Shankman, S. A. (2021). Unique longitudinal relationships between symptoms of psychopathology in youth: A cross-lagged panel network analysis in the ABCD study. *Journal of Child Psychology and Psychiatry*, 62(2), 184–194.

Hamaker, E. L., Kuiper, R. M., & Grasman, R. P. (2015). A critique of the cross-lagged panel model. *Psychological Methods*, 20(1), 102–116.

Hoekstra, R. H., Epskamp, S., Finnemann, A. T., & Borsboom, D. (2023). Unlocking the potential of simulation studies in network psychometrics: A tutorial. <https://doi.org/10.31234/osf.io/4j3hf>

Isvoranu, A.-M., & Epskamp, S. (2022a). Constructing and drawing networks in qgraph. In *Network psychometrics with R* (pp. 67–78). Routledge.

Isvoranu, A.-M., & Epskamp, S. (2022b). Which estimation method to choose in network psychometrics? Deriving guidelines for applied researchers. *Psychological Methods*, 28, 925–946.

Isvoranu, A.-M., Epskamp, S., Waldorp, L., & Borsboom, D. (2022). *Network psychometrics with R: A guide for behavioral and social scientists*. Routledge.

Koelen, J. A., Mansueto, A. C., Finnemann, A., de Koning, L., van der Heijde, C. M., Vonk, P., Wolters, N. E., Klein, A., Epskamp, S., & Wiers, R. W. (2022). COVID-19 and mental health among at-risk university students: A prospective study into risk and protective factors. *International Journal of Methods in Psychiatric Research*, 31(1), e1901.

Liu, A., Ren, Y., Yang, S., Li, Z., Zhu, Z., Zhang, L. M., & Peng, Y. (2024). Contemporaneous and temporal network analysis of complex posttraumatic stress disorder among Chinese college students with childhood adversity: A longitudinal study. *Psychiatry Research*, 336, 115872.

O'Driscoll, C., Epskamp, S., Fried, E. I., Saunders, R., Cardoso, A., Stott, J., Wheatley, J., Cirkovic, M., Naqvi, S. A., Buckman, J. E. J., et al. (2022). Transdiagnostic symptom dynamics during psychotherapy. *Scientific Reports*, 12(1), 10881.

Schade, H. M., Digutsch, J., Kleinsorge, T., & Fan, Y. (2021). Having to work from home: Basic needs, well-being, and motivation. *International Journal of Environmental Research and Public Health*, 18(10), 5149.

Zainal, N. H., & Newman, M. G. (2023a). A cross-lagged prospective network analysis of depression and anxiety and cognitive functioning components in midlife community adult women. *Psychological Medicine*, 53(9), 4160–4171.

Zainal, N. H., & Newman, M. G. (2023b). Prospective network analysis of proinflammatory proteins, lipid markers, and depression components in midlife community women. *Psychological Medicine*, 53(11), 5267–5278.

Reviewer 2:

Review Statement

As a reviewer, I have expertise in network analysis and psychometrics domains. Based on my experience and the content of the manuscript, my review aims to be comprehensive. My review is ordered by section as they appear in the manuscript. My comments are labeled Edit, Minor, and Major for what I see as revisions to the manuscript. The editor, other reviewers, and authors may disagree.

Review

Minor: On page 7, the authors state, “Conditional association networks can be estimated using l_1 -regularized regression techniques, which push many edges to exactly zero, resulting in the ‘sparsest’ network that can still account for associations in the data.” A bit pedantic but I’m not sure that the “sparsest” network is true as the model could be sparser (indeed, non-regularized approaches often end up sparser than l_1 -regularized methods. I believe “sparse” can suffice to the authors’ point.

Minor: In the footnote on page 7, it’s also worth mentioned the {mgm}, {BGGM}, and {GGMnonreg} packages to be comprehensive on packages that can estimate partial correlation networks.

Minor: On page 9 the authors mention the stationary assumption of (vector) autoregressive models, which is often a personal concern I have over AR models. Later on page 11, the authors mention that, “A key benefit of this approach is that we do not need the assumption of stationarity to estimate the model, and, as such, can identify when the patterns of effects vary across timepoints.” Can the authors go into more detail as to why? It’s not clear to me from the text and would be important for readers to understand why they can disregard the assumption despite the use of autoregression.

Minor: On pages 13-14, the authors discussing collapsing nodes based mainly on semantic similarity. Is the idea that researchers should use their discretion to make these decisions? This section lacks a clear sense of direction for the applied researcher. They mention one statistical method in the footnote and another statistical method could be looking at variation inflation factors given the models are initially estimated with regression. Other options might include natural language

processing, which seems to jibe most closely to the idea of semantic similarity, by using a method such as [sentence similarity](#) or word embeddings (Wulff & Mata, 2023).

Wulff, D. U., & Mata, R. (2023). Automated jingle–jangle detection: Using embeddings to tackle taxonomic incommensurability. PsyArXiv.
<https://doi.org/10.31234/osf.io/9h7aw>

Minor: When describing the select of λ on page 16, I'm unclear how it is actually select. The authors state, "Though there are many criteria that can be applied to choose λ_j , here we use 10-fold cross-validation: Regression estimates are obtained for a sequence of 100 λ values, and the one that produces the lowest cross-validation error is chosen."

With 10-fold cross-validation, there will be 10 λ values that are produced (each fold producing a λ from the sequence of 100). Is the mean taken in this case? Or is it the case the there are 100 λ values that are tested across the folds and the one that produces the lowest average (I assume EBIC) across the folds is chosen?

Minor: On the interpretation of individual paths, I would be interested in knowing how many students in the sample might fall into their interpretations that are provided; for example, "While these negative paths appear to contradict the general positive associations among commitment to school variables (e.g., finding school interesting strongly positively predicts and is predicted by enjoyment of school), they might reflect a tendency for students who see themselves as having mastered the material to feel bored in class." (p. 29).

This interpretation is plausibly true but is vague in actionable terms. Can we take this interpretation as general to the sample? I recognize this question gets to the point of between-person and within-person explanations that may not be able to be entertained with the present model; however, might there be some way of identifying this trajectory for individual students (or even groups of students)? (Note: The authors bring up the point of ergodicity on page 37-38 which is particularly relevant to these interpretations)

Minor: Perhaps a point on the predictive power of variable-level over aggregate-level could point to similar findings in personality where the unique variance of items often enhance predictive power above and beyond their corresponding aggregates (e.g., facets, traits; Seeboth & Möttus, 2018).

Seeboth, A., & Möttus, R. (2018). Successful explanations start with accurate descriptions: Questionnaire items as personality markers for more accurate predictions. *European Journal of Personality*, 32(3), 186-201.
<https://doi.org/10.1002/per.2147>

Minor: On the RI-CLPM mentioned by the authors on page (38), the authors state that a latent factor to each variable in the network. I might be missing something but

couldn't the residual correlation matrix after the random-intercept is taken across all variables (so a single factor model representing the random-intercept). This residual correlation matrix could then be leveraged in the network analysis (e.g., Garcia-Pardina et al., accepted). I might be missing something on how such an approach might not be possible within the current framework.

Garcia-Pardina, A., Abad, F. J., Christensen, A. P., Golino, H., & Garrido, L. E. (accepted). Dimensionality assessment in the presence of wording effects: A network psychometric and factorial approach. *Behavior Research Methods*.
<https://doi.org/10.31234/osf.io/7yqau>

Overall: I found this paper easy to read and well explained. The authors are introducing a method that I know firsthand has impact — several folks have already asked me how to implement it — so it was a pleasure to review this paper. I believe there are additional points that can be expanded and clarified but otherwise feel that this manuscript is polished.

Signed,

Alexander Christensen

alexpaulchristensen@gmail.com